ISO
TC184/SC4

SC4 Data Architecture - Mapping and Integration Methodology

Outline for presentation and discussion at the WG10 workshop at Shell Centre, London, 1999-04-21/23

Julian Fowler, PDT Solutions

WG10 N256 (update to N253)

ISO
TC184/SC4

Introduction

- This is an updated version of N253 (1999-04-12)
- Changes/additions:
  - more detail on procedures
  - registration and maintenance process
- More detailed discussion is available in paper N255 (1999-04-21)

ISO
TC184/SC4

SC4 Data Architecture - Mapping & Integration Methodology

- Data Architecture PWI deliverables:
  - Architecture overview (N254)
  - Integration model (N220)
  - **Methodology (N255/N256)**
  - Language requirements
    - EXPRESS (N249)
    - EXIST (N2xx)
  - Demonstration

ISO
TC184/SC4

This document ...

- … presents an outline for the contents of a formal description of the Methodology component of the SC4 Data Architecture
- Development of even an initial draft is limited by funding constraints ☹

ISO
TC184/SC4

The final document ...

- Procedures
  - "What to do"
- Practices - analysis
  - "How to do it"
- Practices - languages and tools
  - "How are languages & tools applied"
- Guidelines
  - "What's the best way to do this?"

ISO
TC184/SC4

Mapping and integration methodology - requirements

- Methods describing the following processes:
  - extending the integration model to meet new requirements
  - selecting a "subset" of the integration model that satisfies the semantics of a particular external/application model
  - defining the mapping(s) between the selected subset and the structure of the external/application model

ISO
TC184/SC4

Fundamental concepts and assumptions

- The SC4 data architecture has the form described in WG10 N254
- The "core" of the data architecture is an integration model that consists of:
  - a generic data model
  - reference data (instances of the generic model)
- The integration model has a universal context
  - it is neutral with respect to all external views

ISO
TC184/SC4

Fundamental concepts and assumptions (continued)

- The scope of the integration model is the "union" of the scopes of all models that have been integrated with it.
- Key difference between the IM and STEP's IRs:
  - IRs contain generalizations of the semantics of the models that use it (AIMs)
  - IM contains *precisely* the semantics of all models that use it, by definition
  - In the IM, "interpretation" = "subsetting"

ISO
TC184/SC4

Integration method

- Objective: extend the IM to include concepts not previously explicit within its context and scope
- Two types of extension
  - *Context* extension - requirements are discovered for data that is not covered *at all* by the Integration Model
  - If the premises of the Integration Model are correct, this scenario is unlikely
  - See Wenzel (N147) - "IM Extensions"

ISO
TC184/SC4

Integration method (continued)

- Second type of extension:
  - *scope extension*
  - IM does not include constructs that precisely match the required semantics …
  - … but can support them by refinement (specialization and/or instantiation), or by combination of what is already there

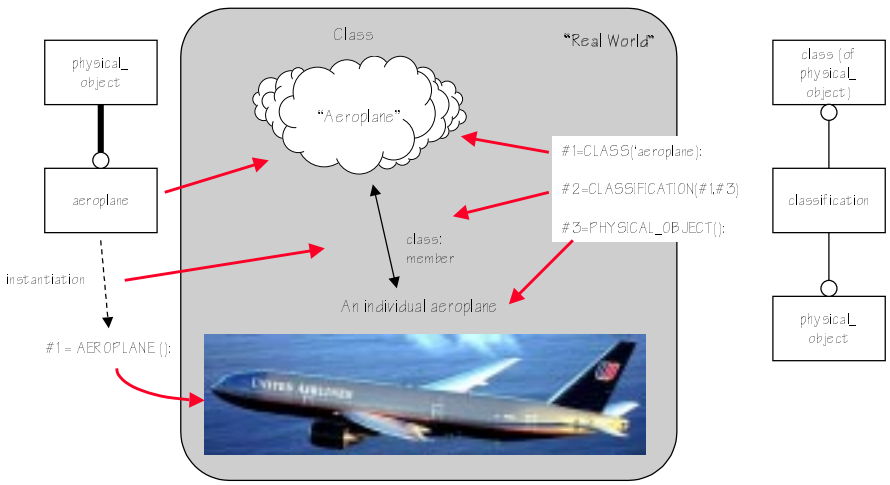
ISO
TC184/SC4

Key issue for the integration method (as yet unanswered)

- When is the model extended by specialization, and when is it extended by instantiation?
- Example:
  - IM (N220) contains entity data type **physical_object**
  - Requirement is for aeroplane (class) - is this:
    - a subtype of **physical_object**, or
    - an instance of **class (of physical_object)**

ISO
TC184/SC4

Specialization vs. Instantiation



ISO
TC184/SC4

Proposed axiom

- The integration model contains as data model elements (EXPRESS entity data types) a minimal set of concepts using which all other facts can be expressed as data
  - Wenzel "Semantically irreducible" elements
- All scope extensions (integration) is achieved through population of the IM structures

ISO
TC184/SC4

Consequences

- If we use EXPRESS for the IM structure, we have to use something like Part 21 or EXPRESS-I to represent populations of the IM
  - may not be natural for modellers!
  - does EXIST help? what about UML? (see Friday discussions)
- Any population of the IM can be expressed (projected) as a model structure

ISO
TC184/SC4

Proposed axiom (2)

- Context extensions are addressed in the structure of the IM

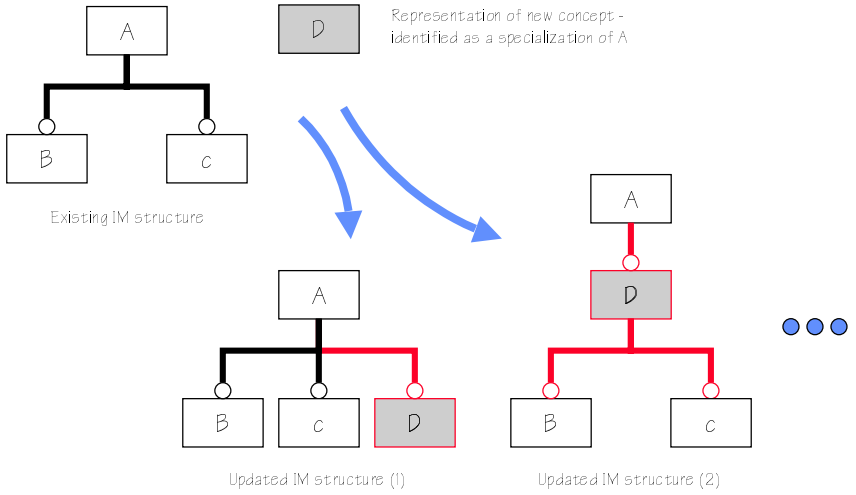
ISO
TC184/SC4

Context extension

- Method for context extension
  - adding a new *irreducible concept* to the IM
- Procedure
  - identify the most specific construct in the existing IM that corresponds to the semantics of the extension
- Practice
  - add a new subtype within the structure of the IM
  - consider relationships to existing subtypes of "parent" (sibling or new intermediate)

ISO
TC184/SC4

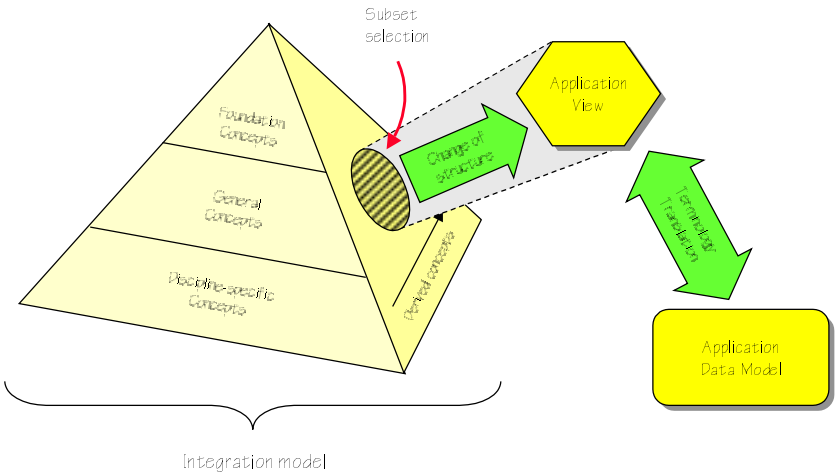
Analysis of sibling relationships



1999-04-22 ISO TC184/SC4 WG10 N256 slide 17

ISO
TC184/SC4

Integration Architecture



1999-04-22 ISO TC184/SC4 WG10 N256 slide 18

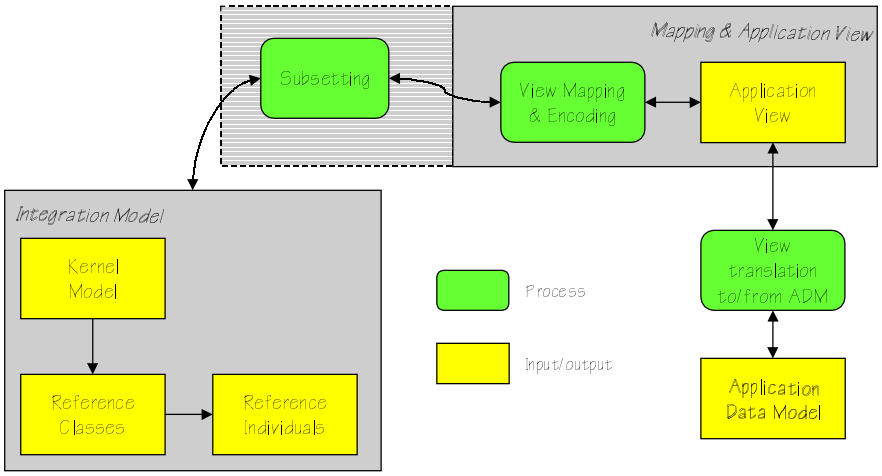
ISO
TC184/SC4

Mapping

- Specification of the transformations between a "subset" of the IM and an external/application data model
- Includes:
  - structural changes
  - terminology changes
- Exclude semantic changes
- I.e., *precisely* the relationship between ARM and AIM in a STEP AP

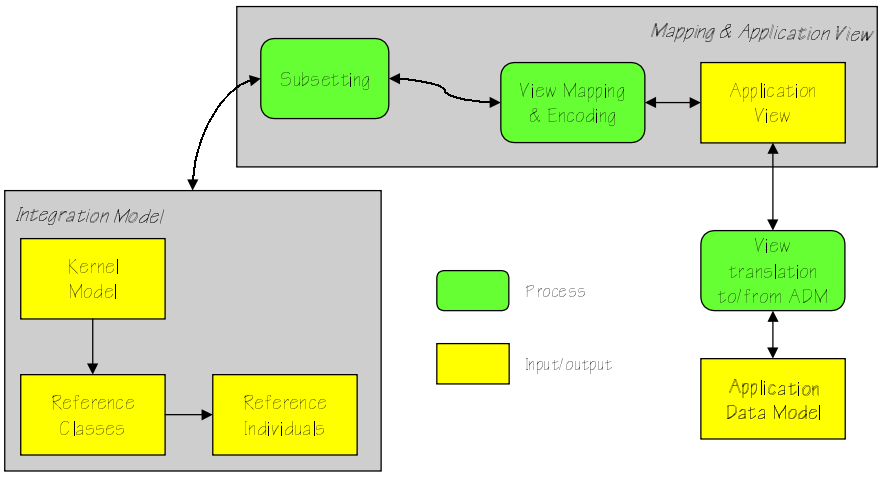
ISO
TC184/SC4

Organization of Integration Architecture



ISO
TC184/SC4

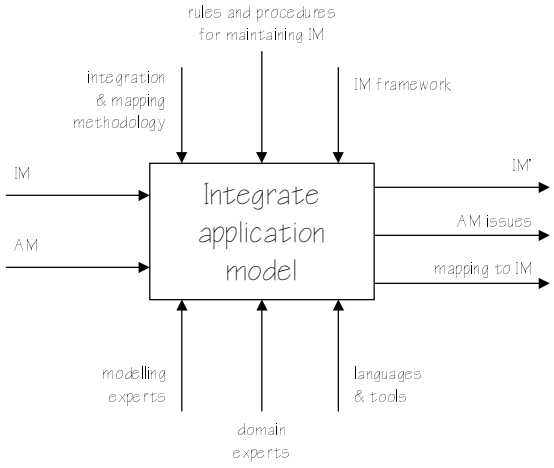
Organization of Integration Architecture



1999-04-22 ISO TC184/SC4 WG10 N256 slide 21

ISO
TC184/SC4

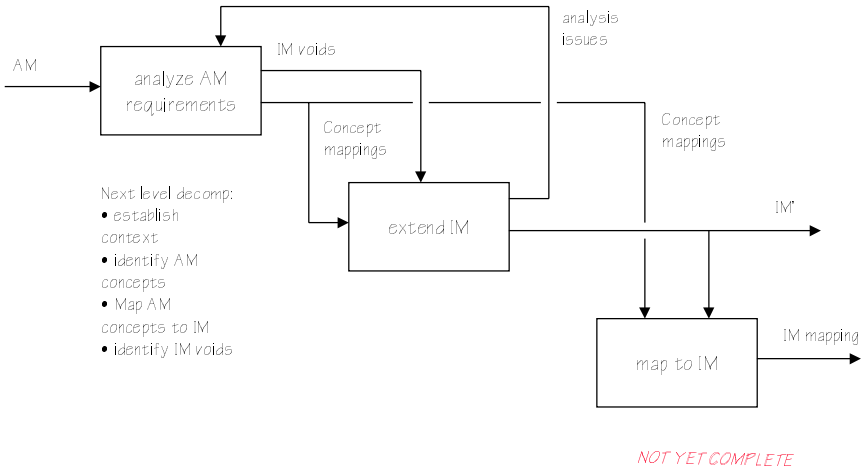
Process overview (A-0)



1999-04-22 ISO TC184/SC4 WG10 N256 slide 22

ISO TC184/SC4

Process overview (A0)



ISO TC184/SC4

Establishing context

- Part of the association between the IM and a particular AM
- "Context" is not really the right word to use here … although this is the same as application *context* in STEP
- Distinguish (for an AM):
  - data scope
  - activity scope ← this is "context"
- Link between data and activity models

ISO
TC184/SC4

More about activity scope

- Options:
- STEP AP approach - describe scope in terms of AAM, identify through population of application_context_schema elements
- Associations to (class of) activity
- Issue for data architecture (not methodology):
  - If activity scope is to be captured, does this go in the IM itself, or as part of standard "views" on the IM

ISO
TC184/SC4

Documenting mappings

- Issue:
  - Document IM ↔ AM mapping in one STEP
  - Document separately:
    - IM subset
    - Model transformation(s)
    - Terminology
- STEP AP paradigm is a mixture of the two (IR - AIM - ARM)
- Best choice: based on reusability

ISO
TC184/SC4

Documenting mappings (continued)

- Recommendation:
  - separate IM ↔ AM mapping into components and document each one
- IM → IM subset
  - is EXPRESS USE FROM sufficient?
- IM subset → AM
  - model transformation (several types)
  - terminology change

ISO
TC184/SC4

Documenting mappings - issues

- N254 assumes separate mappings:
  - projection/encoding (structure)
  - IM subset → Application View
  - terminology change
  - Application View → Application Data Model
  - This needs to be reviewed/discussed - doesn't seem to be consistent (is the AV in this case useful?)
- Language for mappings:
  - EXPRESS-X, MT, others?

ISO
TC184/SC4

Registration - Background

- The Integration Model will be subject to change
  - context extensions (infrequent?)
  - scope extensions (frequent, or at least as often as another model is integrated)
- The frequency of change will not be manageable via "usual" ISO standardization processes
- Proposal: use the ISO *Registration Process* to maintain a *Register* whose content is the Integration Model
  - Foundation Model, Reference Classes, Reference Individuals

ISO
TC184/SC4

ISO/IEC Terminology

- Validation Agency (VA)
  - Agency which receives applications to validate new elements fitting into a set *defined in an International Standard*. Criteria for validation are always defined in the standard.
- Registration Agency (RA)
  - Agency which receives applications to register new elements fitting into a set defined in an International Standard.
- Maintenance Agency (MA)
  - Agency which maintains an International Standard and the elements defined according to it.

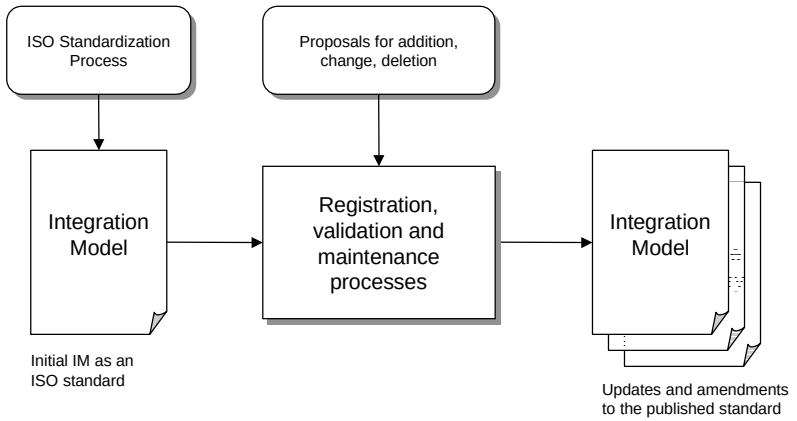
ISO
TC184/SC4

ISO/IEC Terminology (contd.)

- Technical Standard
  - The standard containing the definition of the classes of objects requiring registration
- Procedure Standard
  - The standard containing the specific procedures for the … Registration Agency to follow.
- Establishment of an ISO Register requires:
  - Initial content of the Register, approved as an International Standard
  - Technical Standard
  - Procedure Standard

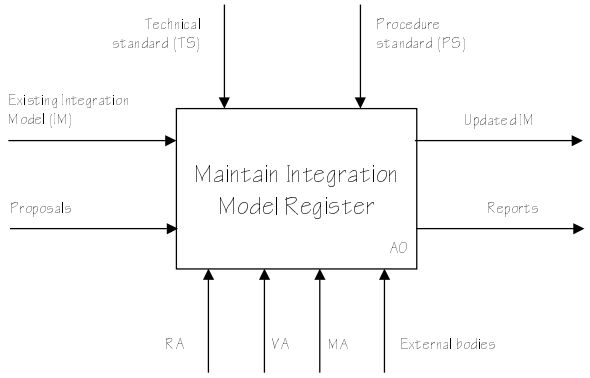
ISO
TC184/SC4

Overview



ISO
TC184/SC4

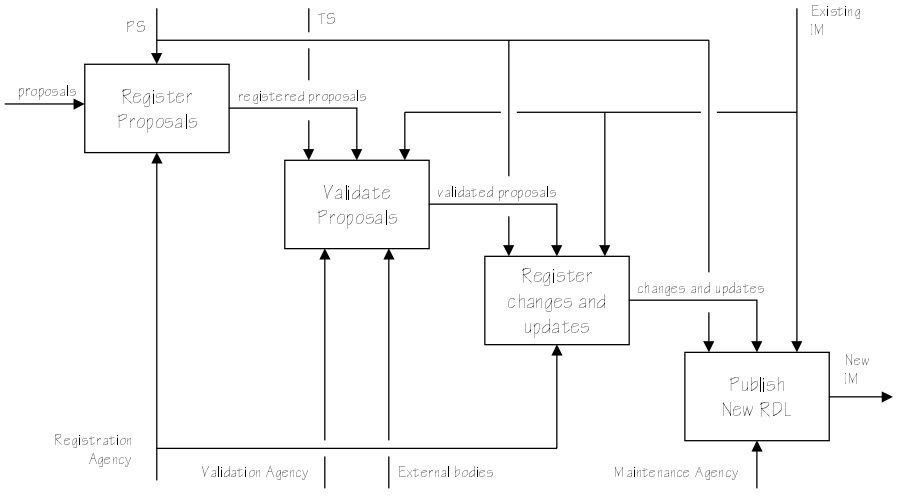
Registration Process



1999-04-22 ISO TC184/SC4 WG10 N256 slide 33

ISO
TC184/SC4

Registration Process



1999-04-22 ISO TC184/SC4 WG10 N256 slide 34

ISO
TC184/SC4

Other relevant standards

- ISO/IEC 11179 "Information technology - Specification and standardization of data elements"
  - Part 6: Registration of data elements
- ISO/IEC JTC1 Directives
  - Defines basic requirements for RA procedures

ISO
TC184/SC4

JTC1 Directives: RA Requirements

- Criteria for applications for registration
- Information to be included on application
- Steps involved in review and response application
- Criteria for rejection of applications
- Procedures for validation (Technical Standard)
- Procedures for maintenance of register
- Requirements for confidentially, protection of IPR, … (if applicable)
- Procedures for publication of the register
- Consultation with other groups
- Dispute resolution
- Others as necessary for the domain ...

ISO
TC184/SC4

Next steps

- April workshop
  - present/discuss ideas
  - experimental mappings
- Lillehammer meeting
  - more substantial outline will be available
  - more detail
  - feedback from London meeting
  - need to schedule a time for discussion